

Nonconvex notions of regularity and convergence of fundamental algorithms for feasibility problems

Robert Hesse* and D. Russell Luke†

December 12, 2012

Abstract

We consider projection algorithms for solving (nonconvex) feasibility problems in Euclidean spaces. Of special interest are the Method of Alternating Projections (MAP) and the Douglas-Rachford or Averaged Alternating Reflection Algorithm (AAR). In the case of convex feasibility, firm nonexpansiveness of projection mappings is a global property that yields global convergence of MAP and for consistent problems AAR. Based on (ϵ, δ) -regularity of sets developed by Bauschke, Luke, Phan and Wang in 2012, a relaxed local version of firm nonexpansiveness with respect to the intersection is introduced for consistent feasibility problems. Together with a coercivity condition that relates to the regularity of the intersection, this yields local linear convergence of MAP for a wide class of nonconvex problems, and even local linear convergence of AAR in more limited nonconvex settings.

2010 Mathematics Subject Classification: Primary 49J52, 49M20; Secondary 47H09, 65K05, 65K10, 90C26.

Keywords: Constraint qualification, linear convergence, method of alternating projections, Douglas Rachford, normal cone, projection operator, reflection operator, set-valued mapping, regularity, metric regularity, strong regularity.

1 Introduction

In the last decade there has been significant progress in the understanding of convergence of algorithms for solving generalized equations, and in particular those arising from variational problems such as minimization or maximization of functions, feasibility, variational inequalities and minimax problems. Early efforts focused on the proximal point algorithm [20, 27] and notions of *hypomonotonicity* which is closely related to *prox-regularity* of functions [28]. Other works have focused on *metric regularity* and its refinements [2, 3, 21]. Proximal-type algorithms have been studied using

*Institut für Numerische und Angewandte Mathematik Universität Göttingen, Lotzestr. 16–18, 37083 Göttingen, Germany. E-mail: hesse@math.uni-goettingen.de.

†Institut für Numerische und Angewandte Mathematik Universität Göttingen, Lotzestr. 16–18, 37083 Göttingen, Germany. E-mail: r.luke@math.uni-goettingen.de.

very different techniques in [4]. In a more limited context, the method of alternating projections (MAP) has been investigated in [9, 24] with the aim of formulating *dual* characterizations of regularity requirements for linear convergence via the *normal cone* and its variants.

The framework we present here generalizes the tools for the analysis of fixed-point iterations of operators that violate the classical property of firm-nonexpansiveness in some quantifiable fashion. As such, our approach is more closely related to the ideas of [27] and hypomonotonicity through the resolvent mapping, however our application to MAP bears a direct resemblance to the more classical *primal* characterizations of regularity described in [7] (in particular what they call “linear regularity”). There are also some parallels between what we call (S, ε) -firm nonexpansiveness and ε -Enlargements of maximal monotone operators [11], though this is beyond the scope of this paper. Our goal is to introduce the essential tools we make use of with a cursory treatment of the connections to other concepts in the literature, and to apply these tools to the MAP and AAR algorithms, comparing our results to the best known results at this time.

We review the basic definitions and classical definitions and results below. In section 2 we introduce our relaxations of the notions of firm-nonexpansiveness and set regularity. Our main abstract result concerning fixed-point iterations of mappings that violate the classical firm-nonexpansive assumptions is in section 3. We specialize in subsequent subsections to MAP and the Douglas-Rachford algorithm, which we refer to as averaged alternating reflections (AAR). Our statement of linear convergence of MAP is as general as the results reported in [9], with more elementary proofs, although our estimates for the radius of convergence are more conservative. The results on local linear convergence of nonconvex instances of AAR to our knowledge are new and provide some evidence supporting the conjecture that, asymptotically, AAR converges more slowly (albeit still linearly) than simple alternating projections. Our estimates of the rate of convergence for both MAP and AAR are not optimal, and so a comparison between these two algorithms is not complete. We show, however, that strong regularity conditions on the set of fixed points are in fact *necessary* for local linear convergence of the AAR algorithm, in contrast to MAP where the same conditions are sufficient, but not necessary [8, 9].

1.1 Basics/Notation

\mathbf{E} is a Euclidean space. We denote the unit ball centered at the origin by \mathbb{B} and the ball of radius δ centered at $\bar{x} \in \mathbf{E}$ by $\mathbb{B}_\delta(\bar{x})$. When the δ -ball is centered at the origin we write \mathbb{B}_δ . The Notation “ \rightrightarrows ” indicates that this mapping in general is *multi-valued*.

Definition 1. Let $\Omega \subset \mathbf{E}$ be nonempty, $x \in \mathbf{E}$. The distance of x to Ω is defined by

$$d(x, \Omega) := \min_{y \in \Omega} \|x - y\|. \quad (1.1)$$

Definition 2 (projectors/reflectors). Let $\Omega \subset \mathbf{E}$ be nonempty and $x \in \mathbf{E}$. The (possibly empty) set of all best approximation points from x to Ω denoted $P_\Omega(x)$ (or $P_\Omega x$), is given by

$$P_\Omega(x) := \{y \in \Omega \mid \|x - y\| = d(x, \Omega)\}. \quad (1.2)$$

The mapping $P_\Omega \rightrightarrows \Omega$ is called the metric projector, or projector, onto Ω . We call an element of this set a projection. The reflector $R_\Omega : \mathbf{E} \rightrightarrows \mathbf{E}$ to the set Ω is defined as

$$R_\Omega x := 2P_\Omega x - x, \quad (1.3)$$

for all $x \in \mathbf{E}$.

Since we are on a Euclidean space \mathbf{E} convexity and closedness of a subset $C \subset \mathbf{E}$ is sufficient for the projector (respectively the reflector) to be single valued. Closedness of a set Ω suffices for the set Ω being *proximal*, i.e. $P_C x \neq \emptyset$ for all $x \in \mathbf{E}$.

Definition 3 (Method of Alternating Projections). For two sets $A, B \subset \mathbf{E}$ we call the mapping

$$T_{MAP}x = P_A P_B x \quad (1.4)$$

the Method of Alternating Projections operator. We call the MAP algorithm, or simply MAP, the corresponding Picard iteration,

$$x_{n+1} \in T_{MAP}x_n, \quad n = 0, 1, 2, \dots \quad (1.5)$$

for x_0 given.

Definition 4 (Averaged Alternating Reflections/Douglas Rachford). For two sets $A, B \subset \mathbf{E}$ we call the mapping

$$T_{AAR}x = \frac{1}{2}(R_A R_B x + x) \quad (1.6)$$

the Averaged Alternating Reflection (AAR) operator. We call the AAR algorithm, or simply AAR, the corresponding Picard iteration,

$$x_{n+1} \in T_{AAR}x_n, \quad n = 0, 1, 2, \dots \quad (1.7)$$

for x_0 given.

Example 5. The following easy examples will appear throughout this work and serve to illustrate the regularity concepts we introduce and the convergence behavior of the algorithms under consideration.

(i) Two lines in \mathbb{R}^2 :

$$\begin{aligned} A &= \{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 = 0\} \subset \mathbb{R}^2 \\ B &= \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 = x_2\} \subset \mathbb{R}^2. \end{aligned}$$

We will see that MAP and AAR converge with a linear rate to the intersection.

(ii) Two lines in \mathbb{R}^3 :

$$\begin{aligned} A &= \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_2 = 0, x_3 = 0\} \subset \mathbb{R}^3 \\ B &= \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 = x_2, x_3 = 0\} \subset \mathbb{R}^3. \end{aligned}$$

After the first iteration step MAP shows exactly the same convergence behavior as in the first example. AAR does not converge. All iterates stay in a plane parallel to $A + B$.

(iii) A line and a ball intersecting in one point:

$$A = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 = 0\} \subset \mathbb{R}^2$$

$$B = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 + (x_2 - 1)^2 \leq 1\}.$$

MAP converges to the intersection, but convergence is arbitrarily slow, the slower the iterates are to the intersection. AAR has fixed points that lie outside the intersection.

(iv) A line and a non-calm set:

$$A = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 = 0\}$$

$$B = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 = x_1 \sin(1/x_1)\}.$$

(v) A circle and a line:

$$A = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 = \sqrt{2}/2\} \subset \mathbb{R}^2$$

$$B = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 + x_2^2 = 1\}.$$

This example is of our particular interest, since it is a simple model case of the *phase retrieval problem*. So far the only *direct* nonconvex convergence results for AAR are related to this model case, see [1, 10]. Local convergence of MAP is covered by [9, 24] as well as by the results in this work.

□

1.2 Classical results for (firmly) nonexpansive mappings

To begin, we recall (firmly) nonexpansive mappings and their natural association with projectors and reflectors on convex sets. We later extend this notion to nonconvex settings where the algorithms involve set-valued mappings.

Definition 6. Let $\Omega \subset \mathbf{E}$ be nonempty. $T : \Omega \rightarrow \mathbf{E}$ is called nonexpansive, if

$$\|Tx - Ty\| \leq \|x - y\| \tag{1.8}$$

holds for all $x, y \in \Omega$.

$T : \Omega \rightarrow \mathbf{E}$ is called firmly nonexpansive, if

$$\|Tx - Ty\|^2 + \|(\text{Id} - T)x - (\text{Id} - T)y\|^2 \leq \|x - y\|^2 \tag{1.9}$$

holds for all $x, y \in \Omega$.

Lemma 7 (Proposition 4.2 [5]). Let $\Omega \subset \mathbf{E}$ be nonempty and let $T : \Omega \rightarrow \mathbf{E}$. The following are equivalent

- (i) T is firmly nonexpansive on Ω
- (ii) $2T - \text{Id}$ is nonexpansive on Ω
- (iii) $\|Tx - Ty\|^2 \leq \langle Tx - Ty, x - y \rangle$ for all $x, y \in \Omega$

Theorem 8 (best approximation property - convex case). *Let $C \subset \mathbf{E}$ be nonempty and convex, $x \in \mathbf{E}$ and $\bar{x} \in C$. \bar{x} is the best approximation point $\bar{x} = P_C(x)$ if and only if*

$$\langle x - \bar{x}, y - \bar{x} \rangle \leq 0 \quad \text{for all } y \in C. \quad (1.10)$$

If C is a affine subspace then (1.10) holds with equality.

Proof. For (1.10) see Theorem 3.14 of [5], while equality follows from Corollary 3.20 of the same. \square

Theorem 9 ((firm) nonexpansiveness of projectors/reflectors). *Let C be a closed, nonempty and convex set. The projector $P_C : \mathbf{E} \rightarrow \mathbf{E}$ is a firmly nonexpansive mapping and hence the reflector R_C is nonexpansive. If, in addition, C is an affine subspace then following conditions hold.*

(i) *P_C is firmly nonexpansive with equality, i.e.*

$$\|P_C x - P_C y\|^2 + \|(\text{Id} - P_C)x - (\text{Id} - P_C)y\|^2 = \|x - y\|^2, \quad (1.11)$$

for all $x \in \mathbf{E}$.

(ii) *For all $x \in \mathbf{E}$*

$$\|R_C x - c\| = \|x - c\| \quad (1.12)$$

holds for all $c \in C$.

Proof. For the first part of the statement see [12, Theorems 4.1 and 5.5], [16, Chapter 12], [17, Propositions 3.5 and 11.2] and [30, Lemma 1.1]. The well-known refinement for affine subspaces follows by a routine application of the definitions and Theorem 8. \square

2 (S, ε) -firm nonexpansiveness

Up to this point, the results have concerned only convex sets, and hence the projector and related algorithms have all been single-valued. In what follows, we generalize to nonconvex sets and therefore allow multi-valuedness of the projectors.

Lemma 10. *Let $A, B \subset \mathbf{E}$ be nonempty and closed. Let $x \in \mathbf{E}$. For any element $x_+ \in T_{AAR}x$ there is a point $\tilde{x} \in R_A R_B$ such that $x_+ = \frac{1}{2}(\tilde{x} + x)$. Moreover, T_{AAR} satisfies the following properties.*

(i)

$$\|x_+ - y_+\|^2 + \|(x - x_+) - (y - y_+)\|^2 = \frac{1}{2} \|x - y\|^2 + \frac{1}{2} \|\tilde{x} - \tilde{y}\|^2 \quad (2.1)$$

where x and y are elements of \mathbf{E} , x_+ and y_+ are elements of $T_{AAR}x$ and $T_{AAR}y$ respectively, and $\tilde{x} \in R_A R_B x$ and $\tilde{y} \in R_A R_B y$ are the corresponding points satisfying $x_+ = \frac{1}{2}(\tilde{x} + x)$ and $y_+ = \frac{1}{2}(\tilde{y} + y)$.

(ii) *For all $x \in \mathbf{E}$*

$$T_{AAR}x = \{P_A(2z - x) - z + x \mid z \in P_B x\}. \quad (2.2)$$

Proof. By Definition 4

$$\begin{aligned} x_+ &\in T_{AAR}x \\ \iff x_+ &\in \frac{1}{2}(R_AR_Bx + x) \\ \iff 2x_+ - x &\in R_AR_Bx. \end{aligned}$$

Defining $\tilde{x} = 2x_+ - x$ yields $x_+ = \frac{1}{2}(\tilde{x} + x)$, where $\tilde{x} \in R_AR_Bx$.

- (i) For $x_+ \in T_{AAR}x$ (respectively $y_+ \in T_{AAR}y$) choose $\tilde{x} \in R_AR_Bx$ (respectively \tilde{y}) such that $x_+ = (\tilde{x} + x)/2$ (respectively y_+). Then

$$\begin{aligned} &\|x_+ - y_+\|^2 + \|(x - x_+) - (y - y_+)\|^2 \\ &= \left\| \frac{1}{2}\tilde{x} + \frac{1}{2}x - \frac{1}{2}\tilde{y} - \frac{1}{2}y \right\|^2 + \left\| \frac{1}{2}x - \frac{1}{2}\tilde{x} - \frac{1}{2}y + \frac{1}{2}\tilde{y} \right\|^2 \\ &= \frac{1}{2}\|x - y\|^2 + \frac{1}{2}\|\tilde{x} - \tilde{y}\|^2 \\ &\quad + \frac{1}{2}\langle \tilde{x} - \tilde{y}, x - y \rangle - \frac{1}{2}\langle \tilde{x} - \tilde{y}, x - y \rangle \\ &= \frac{1}{2}\|x - y\|^2 + \frac{1}{2}\|\tilde{x} - \tilde{y}\|^2. \end{aligned}$$

- (ii) This follows easily from the definitions:

$$\begin{aligned} T_{AAR}x &= \left\{ \frac{1}{2}(R_Av + x) \mid v \in R_Bx \right\} \\ &= \left\{ \frac{1}{2}(R_A(2z - x) + x) \mid z \in P_Bx \right\}, \quad \text{where } v = 2z - x, \ z \in P_Bx \\ &= \left\{ \frac{1}{2}(2P_A(2z - x) - z + x) \mid z \in P_Bx \right\} \\ &= \{P_A(2z - x) - z + x \mid z \in P_Bx\}. \end{aligned}$$

□

Remark 11. In the case where A and B are convex, then as a consequence of Lemma 10 i) and the fact that the reflector R_Ω onto a convex set Ω is nonexpansive, we recover the well-known fact that T_{AAR} is firmly nonexpansive and (2.1) reduces to

$$\begin{aligned} &\|T_{AAR}x - T_{AAR}y\|^2 + \|(\text{Id} - T_{AAR})x - (\text{Id} - T_{AAR})y\|^2 \\ &= \frac{1}{2}\|x - y\|^2 + \frac{1}{2}\|R_AR_Bx - R_AR_By\|^2, \end{aligned} \tag{2.3}$$

while (2.2) reduces to

$$T_{AAR}x = x + P_AR_Bx - P_Bx. \tag{2.4}$$

We define next an analog to firm nonexpansiveness in the nonconvex case with respect to a set S .

Definition 12 ((S, ϵ) -(firmly-)nonexpansive mappings). *Let D and S be nonempty subsets of \mathbf{E} and let T be a (multi-valued) mapping from D to \mathbf{E} .*

i) T is called (S, ε) -nonexpansive on D if

$$\begin{aligned} \|x_+ - \bar{x}_+\| &\leq \sqrt{1 + \varepsilon} \|x - \bar{x}\| \\ \forall x \in D, \forall \bar{x} \in S, \forall x_+ \in Tx, \forall \bar{x}_+ \in T\bar{x}. \end{aligned} \quad (2.5)$$

If 2.5 holds with $\varepsilon = 0$ then we say that T is S -nonexpansive on D .

ii) T is called (S, ε) -firmly nonexpansive on D if

$$\begin{aligned} \|x_+ - \bar{x}_+\|^2 + \|(x - x_+) - (\bar{x} - \bar{x}_+)\|^2 &\leq (1 + \varepsilon) \|x - \bar{x}\|^2 \\ \forall x \in D, \forall \bar{x} \in S, \forall x_+ \in Tx, \forall \bar{x}_+ \in T\bar{x}. \end{aligned} \quad (2.6)$$

If 2.6 holds with $\varepsilon = 0$ then we say that T is S -firmly nonexpansive on D .

Note that, as with (firmly) nonexpansive mappings, the mapping T need not be a self-mapping from D to itself. In the special case where $S = \text{Fix } T$, mappings satisfying (2.5) are also called *quasi-(firmly)-nonexpansive* [5]. Quasi-nonexpansiveness is a restriction of another well-known concept, Fejér monotonicity, to $\text{Fix } T$. Equation (2.6) is a relaxed version of firm nonexpansiveness (1.9). The aim of this work is to expand the theory for projection methods (and in particular MAP and AAR) to the setting where one (or more) of the sets are nonconvex. The classical (firmly) nonexpansive operator on D is $(D, 0)$ -(firmly) nonexpansive on D .

Analogous to the relation between firmly nonexpansive mappings and averaged mappings (see [5, Chapter 4] and references therein) we have the following relationship between (S, ε) -firmly nonexpansive mappings and their $1/2$ -averaged companion mapping.

Lemma 13 ($1/2$ -averaged mappings). *Let $D, S \subset \mathbf{E}$ be nonempty and $T : D \rightrightarrows \mathbf{E}$. The following are equivalent*

(i) T is (S, ε) -firmly nonexpansive on D .

(ii) There is a mapping \hat{T} that is $(S, 2\varepsilon)$ -nonexpansive on D with

$$Tx = \frac{1}{2} (x + \hat{T}x) \quad \forall x \in D. \quad (2.7)$$

Proof. For $x \in D$ define $\hat{T} : D \rightrightarrows \mathbf{E}$ by $\hat{T}x := (2Tx - x)$ and choose $x_+ \in Tx$. Observe that there is a corresponding $\hat{x} \in \hat{T}x$ such that $x_+ = \frac{1}{2}(x + \hat{x})$. Let $\bar{x} \in S$, $\bar{x}_+ \in T\bar{x}$. Then

$$\begin{aligned} &\|x_+ - \bar{x}_+\|^2 + \|x - x_+ - (\bar{x} - \bar{x}_+)\|^2 \\ &= \left\| \frac{1}{2}(x + \hat{x}) - \frac{1}{2}(\bar{x} + \hat{\bar{x}}) \right\|^2 + \left\| \frac{1}{2}(x - \hat{x}) - \frac{1}{2}(\bar{x} - \hat{\bar{x}}) \right\|^2 \\ &= \frac{1}{4} \left[\|x - \bar{x}\|^2 + \langle x - \bar{x}, \hat{x} - \hat{\bar{x}} \rangle + \|\hat{x} - \hat{\bar{x}}\|^2 \right] \\ &\quad + \frac{1}{4} \left[\|x - \bar{x}\|^2 - \langle x - \bar{x}, \hat{x} - \hat{\bar{x}} \rangle + \|\hat{x} - \hat{\bar{x}}\|^2 \right] \\ &= \frac{1}{2} \|x - \bar{x}\|^2 + \frac{1}{2} \|\hat{x} - \hat{\bar{x}}\|^2 \\ &\stackrel{!}{\leq} \frac{1}{2} \|x - \bar{x}\|^2 + \frac{1}{2} (1 + 2\varepsilon) \|x - \bar{x}\|^2 \\ &= (1 + \varepsilon) \|x - \bar{x}\|^2 \end{aligned}$$

where the inequality holds if and only if \hat{T} is 2ε -quasinonexpansive. \square

2.1 Regularity of Sets

To achieve property (2.6) for the projector and the AAR operator, we have to require some properties to the sets A and B upon which we project. These are developed next.

Definition 14 (Prox-regularity). *A nonempty (locally) closed set $\Omega \subset \mathbf{E}$ is prox-regular at a point $\bar{x} \in \Omega$ if the projector P_Ω is single-valued around \bar{x} .*

What we take as the definition of prox-regularity actually follows from the equivalence of prox-regularity of sets as defined in [28, Definition 1.1] and the single-valuedness of the projection operator on neighborhoods of the set [28, Theorem 1.3].

Definition 15 (normal cones). *The Fréchet normal cone $\hat{N}_\Omega(\bar{x})$ and the limiting normal cone $N_\Omega(\bar{x})$ to a set $\Omega \subset \mathbf{E}$ at a point $\bar{x} \in \Omega$ are*

$$\hat{N}_\Omega(\bar{x}) := \left\{ v \in \mathbf{E} : \limsup_{x \xrightarrow{\Omega} \bar{x}} \frac{\langle v, x - \bar{x} \rangle}{\|x - \bar{x}\|} \leq 0 \right\}, \quad (2.8)$$

$$N_\Omega(\bar{x}) := \limsup_{x \xrightarrow{\Omega} \bar{x}} \hat{N}_\Omega(x). \quad (2.9)$$

The construction of the *limiting normal cone* goes back to Mordukhovich (see [29, Chap. 6 Commentary]).

Proposition 16 (Mordukhovich). *The limiting normal cone or Mordukhovich normal cone is the smallest cone satisfying the two properties*

1. $y \in P_\Omega(x) \Rightarrow x - y \in N_\Omega(y)$,
2. for any sequence $x_i \rightarrow \bar{x}$ in Ω any limit of a sequence of normals $v_i \in N_\Omega(x_i)$ must lie in $N_\Omega(\bar{x})$.

Definition 17 (Clarke regularity). *A nonempty (locally) closed set $\Omega \subset \mathbf{E}$ is Clarke regular at a point $\bar{x} \in \Omega$ if the Fréchet normal cone and the limiting normal cone coincide, i.e. $N_\Omega(\bar{x}) = \hat{N}_\Omega(\bar{x})$.*

Definition 18 (ε, δ) -(sub)regularity).

- i) *A nonempty set $\Omega \subset \mathbf{E}$ is (ε, δ) -subregular at \bar{x} with respect to $S \subset \mathbf{E}$, if there exists $\varepsilon > 0, \delta > 0$, and*

$$\langle v, z - y \rangle \leq \varepsilon \|v\| \|z - y\| \quad (2.10)$$

holds for all $y \in \mathbb{B}_\delta(\bar{x}) \cap \Omega$, $z \in S \cap \mathbb{B}_\delta(\bar{x})$, $v \in N_\Omega(y)$. We simply say Ω is (ε, δ) -subregular at \bar{x} if $S = \{\bar{x}\}$.

- ii) *If $S = \Omega$ in i) then we say that the set Ω is (ε, δ) -regular at \bar{x} .*
- iii) *If for all $\varepsilon > 0$ there exists a $\delta > 0$ such that (2.10) holds for all $y, z \in \mathbb{B}_\delta(\bar{x}) \cap \Omega$ and $v \in N_\Omega(y)$, then Ω is said to be super-regular.*

The definition of (ε, δ) -regularity was introduced in [9, Definition 9.1] and is a generalization of the notion of super-regularity introduced in [24, Definition 4.3]. More details to (ε, δ) -regularity can be seen in [9]. Of particular interest is the following.

Proposition 19 (Prox-regular implies super-regular, Proposition 4.9 of [24]). *If a closed set $\Omega \subset \mathbf{E}$ is prox-regular at a point in Ω , then it is super-regular at that point. If a closed set $\Omega \subset \mathbf{E}$ is super-regular at a point in Ω , then it is Clarke regular at that point.*

Superregularity is something between Clarke regularity and amenability or prox-regularity. (ϵ, δ) -regularity is weaker still than Clarke regularity (and hence super-regularity) as the next example shows.

Remark 20. (ϵ, δ) -regularity does not imply Clarke regularity

Proof.

$$\Omega := \left\{ (x_1, x_2) \in \mathbb{R}^2 \left| \begin{array}{ll} x_2 \leq -x_1 & x_1 \leq 0 \\ x_2 \leq 0 & x_1 > 0 \end{array} \right. \right\}.$$

For any $x_- \in \partial_- \Omega := \{(x_1, x_2) | x_2 = -x_1, x_2 \leq 0\}$ and $x_+ \in \partial_+ \Omega := \{(x_1, 0) | x_1 \geq 0\}$ one has

$$\begin{aligned} N_\Omega(x_-) &= \{(\lambda, \lambda) | \lambda \in \mathbb{R}^+\} \\ N_\Omega(x_+) &= \{(0, \lambda) | \lambda \in \mathbb{R}^+\} \end{aligned}$$

which implies $N_\Omega(0) = N_\Omega(x_-) \cup N_\Omega(x_+)$. Since $\widehat{N}_\Omega(0) = \{0\}$ the set Ω is not Clarke regular at 0. Define $\nu_- = (\sqrt{2}/2, \sqrt{2}/2) \in N_\Omega(x_-)$, $\nu_+ = (0, 1) \in N_\Omega(x_+)$ and note

$$\langle \nu_-, 0 - x_- \rangle = 0 \quad \text{and} \quad \langle \nu_+, 0 - x_+ \rangle = 0$$

to show that Ω is $(0, \infty)$ -subregular at 0. By the use of these two inequalities one now has

$$\begin{aligned} \langle \nu_-, x_+ - x_- \rangle &= \langle \nu_-, x_+ \rangle \leq \sqrt{2}/2 \|x_+\| \\ \langle \nu_+, x_- - x_+ \rangle &= \langle \nu_+, x_- \rangle \leq \sqrt{2}/2 \|x_-\| \end{aligned}$$

and hence Ω is $(\frac{\sqrt{2}}{2}, \infty)$ -regular. \square

2.2 Projectors and Reflectors

We are now ready to apply the above definitions in order to establish (S, ϵ) (firm)-nonexpansiveness of projectors and reflectors.

Theorem 21 (projectors and reflectors onto (ϵ, δ) -subregular sets). *Let $\Omega \subset \mathbf{E}$ be nonempty closed and (ϵ, δ) -subregular at each point $\bar{x} \in S \subset \Omega$.*

(i) *The projector is $(S, \tilde{\epsilon}_1)$ -nonexpansive on $\mathbb{B}_\delta(S)$ where $\tilde{\epsilon}_1 \equiv 2\epsilon + \epsilon^2$: at each $\bar{x} \in S$*

$$\|x_+ - \bar{x}\| \leq \underbrace{\sqrt{1 + \tilde{\epsilon}_1}}_{=(1+\epsilon)} \|x - \bar{x}\| \quad \forall x \in \mathbb{B}_\delta(\bar{x}), x_+ \in P_\Omega x. \quad (2.11)$$

(ii) *The projector is $(S, \tilde{\epsilon}_2)$ -firmly nonexpansive on $\mathbb{B}_\delta(S)$, where $\tilde{\epsilon}_2 \equiv 2\epsilon + 2\epsilon^2$: at each $\bar{x} \in S$*

$$\|x_+ - \bar{x}\|^2 + \|x - x_+\|^2 \leq (1 + \tilde{\epsilon}_2) \|x - \bar{x}\|^2 \quad \forall x \in \mathbb{B}_\delta(\bar{x}), x_+ \in P_\Omega x. \quad (2.12)$$

(iii) *The reflector R_Ω is $(S, \tilde{\epsilon}_3)$ -nonexpansive on $\mathbb{B}_\delta(S)$, where $\tilde{\epsilon}_3 \equiv 4\epsilon + 4\epsilon^2$: at each $\bar{x} \in S$*

$$\|x_+ - \bar{x}\| \leq \sqrt{1 + 4\epsilon(1 + \epsilon)} \|x - \bar{x}\| \quad \forall x \in \mathbb{B}_\delta(\bar{x}), x_+ \in R_\Omega x. \quad (2.13)$$

Proof. (i) The projector is nonempty by closedness. Then by the definition of (ε, δ) -subregularity and the Cauchy-Schwarz inequality, for each $\bar{x} \in S$

$$\begin{aligned}\|x_+ - \bar{x}\|^2 &= \langle x - \bar{x}, x_+ - \bar{x} \rangle + \langle x_+ - x, x_+ - \bar{x} \rangle \\ &\leq \|x - \bar{x}\| \|x_+ - \bar{x}\| + \varepsilon \|x - x_+\| \|x_+ - \bar{x}\| \\ &\leq (1 + \varepsilon) \|x - \bar{x}\| \|x_+ - \bar{x}\| \quad \forall x \in \mathbb{B}_\delta(\bar{x}), x_+ \in P_\Omega x.\end{aligned}$$

Thus at each $\bar{x} \in S$

$$\|x_+ - \bar{x}\| \leq (1 + \varepsilon) \|x - \bar{x}\| = \sqrt{(1 + \varepsilon)^2} \|x - \bar{x}\| = \sqrt{1 + (2\varepsilon + \varepsilon^2)} \|x - \bar{x}\|$$

for all $x \in \mathbb{B}_\delta(\bar{x})$, as claimed.

(ii) By the definition of (ε, δ) -subregularity (2.10), at each $\bar{x} \in S$

$$\begin{aligned}&\|x_+ - \bar{x}\|^2 + \|x - x_+\|^2 \\ &= \|x_+ - \bar{x}\|^2 + \|x - \bar{x} + \bar{x} - x_+\|^2 \\ &= \|x_+ - \bar{x}\|^2 + \|x - \bar{x}\|^2 + 2\langle x - \bar{x}, \bar{x} - x_+ \rangle + \|x_+ - \bar{x}\|^2 \\ &= 2\|x_+ - \bar{x}\|^2 + \|x - \bar{x}\|^2 + 2\underbrace{\langle x_+ - \bar{x}, \bar{x} - x_+ \rangle}_{= -\|x_+ - \bar{x}\|^2} + 2\underbrace{\langle x - x_+, \bar{x} - x_+ \rangle}_{\leq \varepsilon \|x - x_+\| \|x_+ - \bar{x}\|} \\ &\leq \|x - \bar{x}\|^2 + 2\varepsilon \|x_+ - \bar{x}\| \|x - x_+\| \quad \forall x \in \mathbb{B}_\delta(\bar{x}), x_+ \in P_\Omega x.\end{aligned}\tag{2.14}$$

By definition $\|x - x_+\| \leq \|x - \bar{x}\|$. Combining (2.14) and equation (2.11) yields at each $\bar{x} \in S$

$$\|x_+ - \bar{x}\|^2 + \|x - x_+\|^2 \leq (1 + 2\varepsilon(1 + \varepsilon)) \|x - \bar{x}\|^2 \quad \forall x \in \mathbb{B}_\delta(\bar{x}), x_+ \in P_\Omega x.$$

(iii) For each $\bar{x} \in S$ by (ii) the projector is $(\{\bar{x}\}, 2\varepsilon + 2\varepsilon^2)$ -firmly nonexpansive on $\mathbb{B}_\delta(\bar{x})$, and so by Lemma 13 (ii) R_Ω is $(\{\bar{x}\}, 4\varepsilon + 4\varepsilon^2)$ -nonexpansive there. Since $\bar{x} \in S$ is arbitrary this proves the result. \square

Note that $\tilde{\varepsilon}_1 < \tilde{\varepsilon}_2$ ($\varepsilon > 0$) in the above theorem, in other words, the *degree* to which classical firm nonexpansiveness is violated is greater than the degree to which classical nonexpansiveness is violated. This is as one would expect since firm nonexpansiveness is a stronger property than nonexpansiveness.

We can now characterize the degree to which the AAR operator violates firm-non-expansiveness on neighborhoods of (ε, δ) -subregular sets.

Theorem 22 ($(S, \tilde{\varepsilon})$ -firm nonexpansiveness of T_{AAR}). *Let $A, B \subset \mathbf{E}$ be closed and nonempty. Let A and B be (ε_A, δ) - and (ε_B, δ) -subregular respectively at each $\bar{x} \in S \subset A \cap B$. The AAR operator $T_{AAR} : \mathbf{E} \rightrightarrows \mathbf{E}$ defined by (1.6) is $(S, \tilde{\varepsilon})$ -firm nonexpansive on $\mathbb{B}_\delta(S)$ where $\tilde{\varepsilon} = 2\varepsilon_A(1 + \varepsilon_A) + 2\varepsilon_B(1 + \varepsilon_B) + 8\varepsilon_A(1 + \varepsilon_A)\varepsilon_B(1 + \varepsilon_B)$, that is*

$$\text{at each } \bar{x} \in S, \quad \|x_+ - \bar{x}\|^2 + \|x - x_+\|^2 \leq (1 + \tilde{\varepsilon}) \|x - \bar{x}\|^2 \tag{2.15}$$

$$\forall x \in \mathbb{B}_\delta(\bar{x}), x_+ \in T_{AAR}x. \tag{2.16}$$

Proof. Choose any $\bar{x} \in S$. Note that $R_A R_B \bar{x} = R_B \bar{x} = \bar{x}$ since $\bar{x} \in A \cap B$. For $x_+ \in T_{AAR}x$ choose $\tilde{x} \in R_A R_B x$, such that $x_+ = (\tilde{x} + x)/2$. Then by Theorem 21(iii)

we have for all $x \in \mathbb{B}_\delta(\bar{x})$

$$\begin{aligned}\|\tilde{x} - \bar{x}\| &\leq \sqrt{1 + 4\varepsilon_A(1 + \varepsilon_A)} \|y - \bar{x}\|, \quad \text{for some } y \in R_B x \\ &\leq \sqrt{1 + 4\varepsilon_A(1 + \varepsilon_A)} \sqrt{1 + 4\varepsilon_B(1 + \varepsilon_B)} \|x - \bar{x}\| \\ &= \sqrt{1 + 2\tilde{\varepsilon}} \|x - \bar{x}\|.\end{aligned}$$

The result then follows from Lemma 13 (i). \square

If one of the sets above is convex, say B for instance, the constant $\tilde{\varepsilon}$ simplifies to $\tilde{\varepsilon} = 2\varepsilon_A(1 + \varepsilon_A)$ since B is $(0, \infty)$ -subregular at \bar{x} . If both sets are convex the we recover 10 iii).

3 Linear Convergence of Iterated (S, ε) -firmly non-expansive Operators

Our goal in this section is to establish the weakest conditions we can (at the moment) under which the MAP and AAR algorithms converge locally linearly. The notions of regularity developed in the previous section are necessary, but not sufficient. In addition to regularity of the operators, we need regularity of the fixed point sets of the operators. This is developed next.

Despite its simplicity, the following Lemma is one of our fundamental tools.

Lemma 23. *Let $D, S \subset \mathbf{E}$, $T : D \rightrightarrows \mathbf{E}$ and $U \subset D \setminus \text{Fix } T$. If*

- (a) *T is (S, ε) -firmly nonexpansive on $D \subset \mathbf{E}$ and*
- (b) *for some point $\bar{x} \in S \subset \mathbf{E}$*

$$\|x - x_+ - (\bar{x} - \bar{x}_+)\| \geq \lambda \|x - \bar{x}\| \quad \forall \bar{x}_+ \in T\bar{x}, \forall x_+ \in Tx, \forall x \in U, \quad (3.1)$$

then T is $(\{\bar{x}\}, \varepsilon - \lambda^2)$ -nonexpansive on U :

$$\|x_+ - \bar{x}_+\| \leq \sqrt{(1 + \varepsilon - \lambda^2)} \|x - \bar{x}\| \quad \forall \bar{x}_+ \in T\bar{x}, \forall x_+ \in Tx, \forall x \in U. \quad (3.2)$$

In particular, for $S \subset \text{Fix } T$,

$$\|x_+ - \bar{x}\| \leq \sqrt{(1 + \varepsilon - \lambda^2)} \|x - \bar{x}\| \quad \forall x_+ \in Tx, \forall x \in U. \quad (3.3)$$

Proof. Choose any $x_+ \in Tx$ and $\bar{x}_+ \in T\bar{x}$. Combine equations (3.1) and (2.6) to get

$$\begin{aligned}\|x_+ - \bar{x}_+\|^2 + (\lambda \|x - \bar{x}\|)^2 &\leq \\ \|x_+ - \bar{x}_+\|^2 + \|x - x_+ - (\bar{x} - \bar{x}_+)\|^2 &\leq (1 + \varepsilon) \|x - \bar{x}\|^2\end{aligned}$$

for all $x \in U$, which immediately yields (3.2). \square

Based on (ε, δ) -(sub-)regularity it has already been shown that both MAP and AAR are (S, ε) -firmly nonexpansive if $S \subset A \cap B$. What remains to be shown is the existence of a point \bar{x} such that (3.1) is fulfilled.

In what follows we determine conditions on the intersection $A \cap B$ to guarantee that the coercivity condition (3.1) holds. We emphasize that the properties (a) and (b) in Lemma 23 are independent. Example 5 (iii) describes the intersection of two

convex sets, so (a) holds, but as we show below (b) fails. As a result the MAP iterates, for example, converge arbitrarily slowly. On the other hand, in Example 5 (iv) describes two sets whose intersection is “well behaved” even though one of the sets is not regular. Combining both properties (a) and (b) in the end (and having a proper balance between them) yields convergence.

3.1 Regularity of Intersections

We define two notions of regularity of the intersection of sets. The first one, which we call *strong regularity*, has many different names in the literature, among them *linear regularity* [24]. We will use the term *linear regularity* to denote a weaker regularity than strong linearity. Both terms “strong” and “linear” are overused in the literature but we have attempted, at the risk of some confusion, to conform to the usage that best indicates the heritage of the concepts.

Definition 24 (strong regularity, Kruger [22]). *The collection of sets $\Omega_1, \Omega_2, \dots, \Omega_m$ is strongly regular at \bar{x} if there exists an $\alpha > 0$ and a $\delta > 0$ such that*

$$(\cap_{i=1}^m (\Omega_i - \omega_i - a_i)) \cap \mathbb{B}_\rho \neq \emptyset \quad (3.4)$$

for all $\rho \in (0, \delta]$, $\omega_i \in \Omega_i \cap \mathbb{B}_\delta(\bar{x})$, $a_i \in B_{\alpha\rho}$, $i = 1, 2, \dots, m$.

Theorem 25 (strongly regular intersection, Theorem 1 [23]). *A family of m closed, nonempty sets $\Omega_1, \Omega_2, \dots, \Omega_m$ has strongly regular intersection at $\bar{x} \in \cap_{j=1}^m \Omega_j$ if and only if there exists a $\kappa > 0$ and a $\delta > 0$ such that*

$$d(x, \cap_{j=1}^m (\Omega_j - x_j)) \leq \kappa \max_{i=1, \dots, m} d(x + x_i, \Omega_i), \quad \forall x \in \mathbb{B}_\delta(\bar{x}), \quad (3.5)$$

for all $x \in \mathbb{B}_\delta(\bar{x})$, $x_i \in \mathbb{B}_\delta$, $i = 1, \dots, m$.

Theorem 26 (Theorem 1 [22]). *A family of closed sets $\Omega_1, \Omega_2, \dots, \Omega_m \subset \mathbf{E}$ has strongly regular intersection (3.4) at a point $\bar{x} \in \cap_i \Omega_i$, if the only solution to the system*

$$\sum_{i=1}^m v_i = 0, \quad \text{with } v_i \in N_{\Omega_i}(\bar{x}) \quad \text{for } i = 1, 2, \dots, m$$

is $v_i = 0$ for $i = 1, 2, \dots, m$. For two sets $\Omega_1, \Omega_2 \subset \mathbf{E}$ this can be written as

$$N_{\Omega_1}(\bar{x}) \cap -N_{\Omega_2}(\bar{x}) = \{0\}, \quad (3.6)$$

and is equivalent to the previous Definition (3.4).

Definition 27 (linear regularity). *A family of closed, nonempty sets $\Omega_1, \Omega_2, \dots, \Omega_m$ has locally linearly regular intersection at $\bar{x} \in \cap_{j=1}^m \Omega_j$ if there exists a $\kappa > 0$ and a $\delta > 0$ such that*

$$d(x, \cap_{j=1}^m \Omega_j) \leq \kappa \max_{i=1, \dots, m} d(x, \Omega_i), \quad \forall x \in \mathbb{B}_\delta(\bar{x}). \quad (3.7)$$

If (3.7) holds for any $\delta > 0$ the intersection is called *linearly regular*. The infimum over all κ such that (3.7) holds is called *regularity modulus*.

Since (3.7) is (3.5) with $x_j = 0$ for all $j = 1, 2, \dots, m$, it is clear that strong regularity implies local linear regularity and is indeed a much more restrictive notion than local linear regularity.

Remark 28. What we are calling local linear regularity at \bar{x} has appeared in various forms elsewhere. See for instance [19, Proposition 4], [26, Section 3], and [23, Equation (15)]. Compare this to (bounded) linear regularity defined in [7, Definition 5.6]. Also compare this to the *basic constraint qualification for sets* in [25, Definition 3.2] and *strong regularity* of the intersection in [23, Proposition 2], also called *linear regularity* in [24].

Remark 29. Based on strong regularity (more specific characterization (3.6)) Lewis, Luke and Malick proved local linear convergence of MAP in the nonconvex setting, where both sets A, B are closed and one of the sets is *super-regular* [24]. This was refined later in [9]. The proof of convergence that will be given in this work is different from the one used in [9, 24] and more related to the one in [7]. Convergence is achieved using (local) linear regularity (3.7), which is described in [14, Theorem 4.5] as *the precise property equivalent to uniform linear convergence of the CPA* (Cyclic projections algorithms). However the rate of convergence achieved by the use of linear regularity is not optimal, while the one in [9, 24] is in some instances. An adequate description of the relation between the *direct/primal* techniques used here and the *dual* approach used in [9, 24] remains open.

Theorem 30 (linear regularity of intersections of convex cones). *Let $\Omega_1, \Omega_2, \dots, \Omega_m$ be a family of closed, nonempty, convex cones. The following statements are equivalent*

- (i) *The intersection is locally linearly regular at \bar{x}*
- (ii) *The intersection is linearly regular at \bar{x}*

Proof. [7] Proposition 5.9 □

Example 31 (Example 5 revisited.). The intersection in example 5 (i) at 0 is strongly regular ($c = \sqrt{2}/2$) and linearly regular ($\kappa = \sqrt{2}/4$). The same example embedded in a higher-dimensional space is still linearly regular, but loses its strong regularity. This can be seen by shifting one of the sets in example 5 (ii) in x_3 -direction, as this renders the intersection empty. This shows that *linear regularity does not imply strong regularity*. The intersection in example (iii) is neither strongly regular nor linearly regular. In each of the examples 5 (iv) and (v) one of the sets is nonconvex, but the intersections are still well-behaved. Observe that in (iv) shifting any of the sets does not render the intersection empty and so the intersection is strongly (locally) regular. Although this already implies linear regularity in (iv), we state the following to show linear regularity more directly. Note that since

$$B \subset \bar{B} := \left\{ (x_1, x_2) \left| \begin{cases} -x_2 \leq x_1 \leq x_2, & x_1 \geq 0 \\ x_2 \leq x_1 \leq -x_2, & x_1 < 0 \end{cases} \right. \right\}$$

by linear regularity of $A \cap \bar{B}$ linear regularity of $A \cap B$ follows. □

The next definitions, while not necessary for our purposes here, provide a bridge between the above notions of regularity and the approaches of [2, 3, 20, 21, 27].

Definition 32 ((strong) metric (sub)-regularity).

- (i) *The mapping $\Phi : \mathbf{E} \rightrightarrows \mathbf{Y}$ is called metrically regular at \bar{x} for \bar{y} if there is a constant $\kappa > 0$ together with neighborhood U of \bar{x} and V of \bar{y} such that*

$$d(x, \Phi^{-1}(y)) \leq \kappa d(y, \Phi(x)) \text{ for all } (x, y) \in U \times V. \quad (3.8)$$

The regularity modulus $\text{reg}\Phi(\bar{x}|\bar{y})$ is the infimum of those constants $\kappa > 0$ such that (3.8) holds.

- (ii) The mapping $\Phi : \mathbf{E} \rightrightarrows \mathbf{Y}$ is called *metrically subregular* at \bar{x} for \bar{y} if $(\bar{x}, \bar{y}) \in \text{gph}\Phi$ and there is a constant $\kappa > 0$ and neighborhoods U of \bar{x} and V of \bar{y} such that

$$d(x, \Phi^{-1}(\bar{y})) \leq \kappa d(\bar{y}, \Phi(x) \cap V) \text{ for all } x \in U. \quad (3.9)$$

The subregularity modulus $\text{subreg}\Phi(\bar{x}|\bar{y})$ is the infimum of those constants $\kappa > 0$ such that (3.9) holds.

- (iii) The mapping $\Phi : \mathbf{E} \rightrightarrows \mathbf{Y}$ is *strongly metrically subregular* at \bar{x} for \bar{y} if $(\bar{x}, \bar{y}) \in \text{gph}\Phi$ and there is a constant $\gamma > 0$ along with neighborhoods U of \bar{x} and V of \bar{y} such that

$$\|x - \bar{x}\| \leq \kappa d(\bar{y}, \Phi(x) \cap V) \quad \text{for all } x \in U. \quad (3.10)$$

Proposition 33 (Metric regularity and strong regularity). *Consider the set-valued mapping $\Phi : \mathbf{E} \rightrightarrows \mathbf{E}^m$*

$$\Phi(x) = (\Omega_1 - x) \times \cdots \times (\Omega_m - x)$$

We have the following characterizations:

- (i) Φ is *metrically regular* at \bar{x} for 0 if and only if the intersection $\cap_i \Omega_i$ is *strongly regular* at \bar{x} .
- (ii) Φ is *metrically subregular* at \bar{x} for 0 if and only if the intersection $\cap_i \Omega_i$ is *locally linearly regular*.
- (iii) Φ is *strongly metrically subregular* at \bar{x} for 0 if and only if the intersection is *locally linearly regular* and $\{\bar{x}\} = (\cap_i \Omega_i) \cap \mathbb{B}_\delta(\bar{x})$ for all $\delta > 0$ small enough.

Proof. The first statement was proved in [24].

To prove the second statement note that

$$\Phi^{-1}((y_1, \dots, y_m)) = \bigcap_{j=1}^m \{x | y_j \in \Omega_j - x\} = \bigcap_{j=1}^m \Omega_j - y_j.$$

Hence

$$\Phi^{-1}((0, \dots, 0)) = \bigcap_{j=1}^m \Omega_j$$

and $(\bar{x}, 0) \in \text{gph}\Phi$ is equivalent to $\bar{x} \in \bigcap_{j=1}^m \Omega_j$. Also by norm equivalence, the existence of $\mathbb{B}_\delta(\bar{x})$ such that

$$d(x, \Phi^{-1}((0, \dots, 0))) \leq \kappa_1 d(0, \Phi(x) \cap V) \quad \text{for all } x \in U$$

is equivalent to the existence of $\kappa_2 > 0$ (different than the κ_1 above) such that

$$d(x, \Omega) \leq \kappa_2 \max_{j=1, \dots, m} d(x, \Omega_j) \quad \text{for all } x \in \mathbb{B}_\delta(\bar{x}).$$

As this is by definition local linear regularity of $(\Omega_1, \dots, \Omega_m)$ at \bar{x} , the proof is complete.

The third property then follows by basic properties of strong metric subregularity, that is that strong metric subregularity at a point \bar{x} for \bar{y} is equivalent to metric subregularity Φ at \bar{x} for \bar{y} and \bar{x} being a *isolated point* of Φ^{-1} . \square

3.2 Linear Convergence of MAP

In the case of the MAP operator, the connection between local linear regularity of the intersection of sets and the coercivity of the operator with respect to the intersection is natural, as the next result shows.

Proposition 34 (coercivity of the projector). *Let A, B be nonempty and closed subsets of \mathbf{E} and let $x_1 \in A$ and $x_2 \in P_B x_1$. If $S \equiv A \cap B$ is locally linearly regular at $\bar{x} \in P_S x_1$ then*

$$\|x_2 - x_1\| \geq \gamma \|x_1 - \bar{x}\| \quad (3.11)$$

where $\gamma = 1/\kappa$ and κ is the regularity modulus.

Proof. Since $x_1 \in A$ one has

$$\begin{aligned} \|x_2 - x_1\| &= d(x_1, B) \\ &= \max \left\{ d(x_1, B), \underbrace{d(x_1, A)}_{=0} \right\} \\ &\geq \gamma d(x_1, S) \\ &= \gamma \|x_1 - \bar{x}\| \end{aligned}$$

where the inequality follows by Definition 27 (local linear regularity at \bar{x}). \square

Theorem 35 (Projections onto a (ε, δ) -subregular set). *Let A, B be nonempty and closed subsets of \mathbf{E} and let $x_1 \in A$ and $x_2 \in P_B x_1$. If*

- (a) $S \equiv A \cap B$ is locally linearly regular at $\bar{x} \in P_S x_1$ and
- (b) B is (ε, δ) -subregular at \bar{x} with $\delta \geq d(x_1, S)$

then

$$d(x_2, S) \leq \sqrt{1 + \tilde{\varepsilon} - \gamma^2} d(x_1, S), \quad (3.12)$$

where $\gamma = 1/\kappa$ with κ the regularity modulus and $\tilde{\varepsilon} = 2\varepsilon + 2\varepsilon^2$.

Proof. By Proposition 34 we have that

$$\|x_2 - x_1\| \geq \gamma \|x_1 - \bar{x}\|.$$

If $x_1 \in S$ then $x_1 = x_2 \in S$ and the result is trivial. Assume then that $x_1 \notin B$. Since $x_1 \in A$ is arbitrary, the coercivity condition on the intersection in (3.1) is satisfied at least on $A \setminus S \subset \mathbf{E} \setminus \text{Fix } P_B$. Since B is (ε, δ) -subregular at \bar{x} , by Theorem 21 the projector P_B is $(\{\bar{x}\}, 2\varepsilon + 2\varepsilon^2)$ -firmly nonexpansive on $\mathbb{B}_\delta(\bar{x})$. In particular, since $\delta \geq d(x_1, S)$ Lemma 23 then applies to yield

$$\|x_2 - \bar{x}\| \leq \sqrt{1 + \tilde{\varepsilon} - \gamma^2} \|x_1 - \bar{x}\|.$$

Since $\bar{x} \in P_S x_1$ we have $d(x_1, S) = \|x_1 - \bar{x}\|$, and by definition $d(x_2, S) \leq \|x_2 - \bar{x}\|$. This completes the proof. \square

The following corollaries are direct consequences of Theorem 35.

Corollary 36 (Projections onto a convex set [18]). *Let A, B be nonempty and closed subsets of \mathbf{E} and let $x_1 \in A$ and $x_2 \in P_B x_1$. If*

- (a) $S \equiv A \cap B$ is locally linearly regular at $\bar{x} \in P_S x_1$ and
- (b) B is convex

then

$$d(x_2, S) \leq \sqrt{1 - \gamma^2} d(x_1, S), \quad (3.13)$$

where $\gamma = 1/\kappa$ with κ the regularity modulus.

Proof. Since B is convex it is $(0, +\infty)$ -regular hence $\tilde{\varepsilon} = 0$ in Theorem 35. \square

Corollary 37 (linear convergence of MAP). *Let A, B be closed nonempty subsets of \mathbf{E} with locally linearly regular intersection at $\bar{x} \in S := A \cap B$. For any $x_0 \in \mathbb{B}_\delta(\bar{x})$, generate the sequence $\{x_n\}_{n \in \mathbb{N}}$ by*

$$x_{2n+1} \in P_A x_{2n} \text{ and } x_{2n+2} \in P_B x_{2n+1} \quad (\forall n = 0, 1, 2, \dots). \quad (3.14)$$

- (a) If A and B are (ε, δ) -subregular w.r.t. S , then

$$d(x_{2n+2}, S) \leq (1 - \gamma^2 + \varepsilon) d(x_{2n}, S).$$

- (b) If A is (ε, δ) -subregular w.r.t. S and B is convex, then

$$d(x_{2n+2}, S) \leq \sqrt{1 - \gamma^2 + \varepsilon} \sqrt{1 - \gamma^2} d(x_{2n}, S).$$

- (c) If A and B are convex, then

$$d(x_{2n+2}, S) \leq (1 - \gamma^2) d(x_{2n}, S).$$

3.3 Linear Convergence of AAR

We now turn to the AAR algorithm. This algorithm is notoriously difficult to analyze and our results reflect this in considerably more circumscribed conditions than are required for the MAP algorithm. Nevertheless, to our knowledge the following convergence results are the most general to date. The first result gives sufficient conditions for the coercivity condition 23 (b) to hold.

Lemma 38. *Let the closed subsets A, B of \mathbf{E} have locally linearly regular intersection at $\bar{x} \in S \equiv A \cap B$ with regularity modulus $\kappa > 0$ and radius of regularity δ in (3.7). Suppose further that B is a subspace and that for some constant $c \in (0, 1)$ the following condition holds:*

$$x \in \mathbb{B}_\delta(\bar{x}), \ y \in P_B x, \ z \in P_A(2y - x) \text{ and } \begin{matrix} u \in N_A(z) \cap \mathbb{B} \\ v \in N_B(y) \cap \mathbb{B} \end{matrix} \Rightarrow \langle u, v \rangle \geq -c. \quad (3.15)$$

Then T_{AAR} fulfills the coercivity condition

$$\|x - x_+\| \geq \frac{\sqrt{1-c}}{\kappa} d(x, S) \quad \forall x \in U \quad (3.16)$$

where $x_+ \in T_{AAR} x$ and $U \equiv \{x \in \mathbb{B}_\delta(\bar{x}) \mid P_A R_B x \subset \mathbb{B}_\delta(\bar{x})\}$.

Proof. In what follows we will use the notation R_Bx for $2y - x$ with $y \in P_Bx$ which is unambiguous, if a slight abuse of notation, since B is convex. We will use (3.15) to show that for all $x \in \mathbb{B}_\delta(\bar{x})$ fixed with $y \in P_Bx$ and $z \in P_AR_Bx$

$$\begin{aligned} \|x - x_+\|^2 &= \|z - y\|^2 \\ &\geq (1 - c) \left[\|z - R_Bx\|^2 + \|R_Bx - y\|^2 \right] \end{aligned} \quad (3.17)$$

We will then show that, for all $x \in U$ fixed with $y \in P_Bx$ and $z \in P_AR_Bx$

$$\|z - R_Bx\|^2 + \|R_Bx - y\|^2 \geq \frac{1}{\kappa^2} d(R_Bx, S)^2. \quad (3.18)$$

Combining inequalities (3.17) and (3.18) yields

$$\|x - x_+\|^2 \geq \frac{1 - c}{\kappa^2} d(R_Bx, S)^2 \quad \forall x \in U.$$

Let $\tilde{x} \in P_S(R_Bx)$ and note that $d(R_Bx, S) = \|R_Bx - \tilde{x}\|$. Since B is a subspace, by (1.12) one has $d(R_Bx, S) = \|R_Bx - \tilde{x}\| = \|x - \tilde{x}\|$.

Moreover, $\|x - \tilde{x}\| \geq \min_{y \in S} \|x - y\| = d(x, S)$, hence

$$\|x - x_+\| \geq \frac{\sqrt{1 - c}}{\kappa} d(x, S) \quad \forall x \in U$$

as claimed.

What remains is to prove (3.17) and (3.18).

Proof of (3.17). Using Lemma 10 equation (2.2) one has for $x \in \mathbb{B}_\delta(\bar{x})$ fixed with $y \in P_Bx$ and $z \in P_AR_Bx$

$$\begin{aligned} &\|x - x_+\|^2 \\ &= \|z - y\|^2 \\ &= \|z - R_Bx + R_Bx - y\|^2 \\ &= \|z - R_Bx\|^2 + \|R_Bx - y\|^2 + 2 \underbrace{\langle z - R_Bx, R_Bx - y \rangle}_{\substack{\in -N_A(z) \\ = y - x \in -N_B(y)}} \\ &\stackrel{3.15}{\geq} \|z - R_Bx\|^2 + \|R_Bx - y\|^2 - 2c \|z - R_Bx\| \|R_Bx - y\| \\ &= (1 - c) \left[\|z - R_Bx\|^2 + \|R_Bx - y\|^2 \right] \\ &\quad + c \left[\|z - R_Bx\|^2 - 2 \|z - R_Bx\| \|R_Bx - y\| + \|R_Bx - y\|^2 \right] \\ &= (1 - c) \left[\|z - R_Bx\|^2 + \|R_Bx - y\|^2 \right] \\ &\quad + c [\|z - R_Bx\| - \|R_Bx - y\|]^2 \\ &\geq (1 - c) \left[\|z - R_Bx\|^2 + \|R_Bx - y\|^2 \right]. \end{aligned} \quad (3.19)$$

△

Proof of (3.18). First note that if $x \in \mathbb{B}_\delta(\bar{x})$, since B is a subspace, by equation (1.12) $R_Bx \subset \mathbb{B}_\delta(\bar{x})$ and by convexity of $\mathbb{B}_\delta(\bar{x})$ it follows that $y = P_Bx \subset \mathbb{B}_\delta(\bar{x})$

and hence (3.15) is localized to $\mathbb{B}_\delta(\bar{x})$ (to which the dividends of linear regularity of the intersection extend) as long as $P_A R_B x \subset \mathbb{B}_\delta(\bar{x})$, that is, as long as $x \in U$. By definition of the projector $\|R_B x - y\| \geq \|R_B x - P_B(R_B x)\|$. Local linear regularity at \bar{x} with radius δ and modulus κ yields

$$\begin{aligned} \|z - R_B x\|^2 + \|R_B x - y\|^2 &\geq \|z - R_B x\|^2 + \|R_B x - P_B R_B x\|^2 \\ &= d(R_B x, A)^2 + d(R_B x, B)^2 \\ &\stackrel{(3.7)}{\geq} \frac{1}{\kappa^2} d(R_B x, S)^2 \quad \forall R_B x \in \mathbb{B}_\delta(\bar{x}) \end{aligned} \quad (3.20)$$

This completes the proof of (3.18) and the Theorem. \square

Remark 39. In (3.19) the difference $[\|z - R_B x\| - \|R_B x - y\|]$ is completely ignored. On the other hand we throw away either $\|z - R_B x\|$ **or** $\|R_B x - P_B R_B x\|$ in (3.20).

If (3.19) is tight then it is the worst possible result in (3.20), since $\|z - R_B x\| = \|R_B x - P_B R_B x\|$. On the other hand if (3.20) is tight this means $\|z - R_B x\|^2 = 0$ **or** $\|R_B x - P_B R_B x\|^2 = 0$ and this means that (3.19) is not tight. This shows that our estimation can never be tight.

Lemma 38 with the added assumption of (ϵ, δ) -regularity of the nonconvex set yields local linear convergence of the AAR algorithm in this special case.

Theorem 40. *Let the closed subsets A, B of \mathbf{E} have locally linearly regular intersection at $\bar{x} \in S \equiv A \cap B$ with regularity modulus $\kappa > 0$ and radius of regularity δ in (3.7). Suppose further that B is a subspace and that A is (ϵ, δ) -regular at \bar{x} . Assume that for some constant $c \in (0, 1)$ the following condition holds:*

$$\left. \begin{array}{l} z \in A \cap \mathbb{B}_\delta(\bar{x}), \quad u \in N_A(z) \cap \mathbb{B} \\ y \in B \cap \mathbb{B}_\delta(\bar{x}), \quad v \in N_B(y) \cap \mathbb{B} \end{array} \right\} \Rightarrow \langle u, v \rangle \geq -c. \quad (3.21)$$

If $x \in (P_S^{-1} \bar{x}) \cap \mathbb{B}_{\frac{\delta}{1+\epsilon}}(\bar{x})$ then

$$d(x_+, S) \leq \sqrt{1 + \tilde{\epsilon} - \eta} d(x, S) \quad \forall x_+ \in T_{AAR} x \quad (3.22)$$

with $\eta := \frac{(1-c)}{\kappa^2}$ and $\tilde{\epsilon} = 2\epsilon + 2\epsilon^2$.

Proof. As in the proof of (3.18), if $x \in \mathbb{B}_{\frac{\delta}{1+\epsilon}}(\bar{x})$, since B is a subspace $R_B x$ and $P_B x \in \mathbb{B}_{\frac{\delta}{1+\epsilon}}(\bar{x})$. Moreover, since A is (δ, ϵ) -regular at \bar{x} , (2.11) yields

$$\|z - \bar{x}\| \leq (1 + \epsilon) \|R_B x - \bar{x}\| \leq (1 + \epsilon) \left(\frac{\delta}{1 + \epsilon} \right) \quad \forall z \in P_A R_B x, \quad \forall R_B x \in \mathbb{B}_\delta(\bar{x})$$

and therefore $P_A R_B x \subset \mathbb{B}_\delta(\bar{x})$. So the set $U = \mathbb{B}_{\frac{\delta}{1+\epsilon}}(\bar{x})$ in (3.16). In particular, on $(P_S^{-1} \bar{x}) \cap \mathbb{B}_{\frac{\delta}{1+\epsilon}}(\bar{x})$ we have

$$\|x - x_+\| \geq \frac{\sqrt{1-c}}{\kappa} d(x, S) = \frac{\sqrt{1-c}}{\kappa} \|x - \bar{x}\|$$

so the coercivity condition (3.1) is satisfied on $(P_S^{-1} \bar{x}) \cap \mathbb{B}_{\frac{\delta}{1+\epsilon}}(\bar{x})$.

Now since A is (ε, δ) -regular and B is $(0, \infty)$ -regular, by Theorem 22 T_{AAR} is $(\{\bar{x}\}, \tilde{\varepsilon})$ -firmly nonexpansive with $\tilde{\varepsilon} = 2\varepsilon(1 + \varepsilon)$, that is

$$\text{at } \bar{x} \in S, \quad \|x_+ - \bar{x}\|^2 + \|x - x_+\|^2 \leq (1 + \tilde{\varepsilon}) \|x - \bar{x}\|^2 \quad \forall x_+ \in T_{AAR}x, \forall x \in \mathbb{B}_\delta(\bar{x}).$$

Lemma 23 then applies to yield

$$\|x_+ - \bar{x}\| \leq \sqrt{1 + \tilde{\varepsilon} - \eta} \|x - \bar{x}\| \quad \forall x_+ \in T_{AAR}x, \forall x \in (P_S^{-1}\bar{x}) \cap \mathbb{B}_{\frac{\delta}{1+\varepsilon}}(\bar{x})$$

where $\eta \equiv \frac{1-c}{\kappa^2}$. The observation that $d(x_+, S) \leq \|x_+ - \bar{x}\|$ for all $x_+ \in T_{AAR}x$ completes the proof. \square

The next lemma establishes sufficient conditions under which (3.21) holds.

Lemma 41 ([24] Theorem 5.16). *Assume $B \subset \mathbf{E}$ is a subspace and that $A \subset \mathbf{E}$ is closed and super-regular at $\bar{x} \in A \cap B$. If the intersection is strongly linearly regular at \bar{x} , then there is a $\delta > 0$ and a constant $c \in (0, 1)$ such that (3.21) holds.*

Proof. Condition (3.15) can be shown using 3.6. For more details see [24]. \square

We summarize this discussion with the following convergence result for the AAR algorithm in the case of an affine subspace and a super-regular set.

Theorem 42. *Assume $B \subset \mathbf{E}$ is a subspace and that $A \subset \mathbf{E}$ is closed and super-regular at $\bar{x} \in S \equiv A \cap B$. If the intersection is strongly linearly regular at S , then there is a $\delta > 0$ such that, $\frac{(1-c)}{\kappa^2} > 2\varepsilon + 2\varepsilon^2$ and hence*

$$d(x_+, S) \leq \tilde{c} d(x, S) \quad \forall x_+ \in T_{AAR}x, \quad (3.23)$$

with $\tilde{c} = \sqrt{1 + 2\varepsilon + 2\varepsilon^2 - \frac{(1-c)}{\kappa^2}} < 1$ for all $x \in (P_S^{-1}\bar{x}) \cap \mathbb{B}_{\frac{\delta}{1+\varepsilon}}(\bar{x})$.

Example 43. The example Example 5 (v) has been studied by Borwein and coauthors [1, 10] where they achieve global characterizations of convergence with rates. Our work does not directly overlap with [1, 10] since our results are local, and the order of the reflectors is reversed: we must reflect first across the subspace, then reflect across the nonconvex set; Borwein and coauthors reflect first across the circle.

3.4 AAR on Subspaces

We finish this section with the surprising fact that strong regularity of the intersection is *necessary*, not just sufficient for convergence of the AAR algorithm.

Corollary 44. *Let A, B be two affine subspaces with $A \cap B$. AAR converges to a point in $A \cap B$ for all $x_0 \in \mathbf{E}$ with linear rate $\tilde{c} < 1$ if and only if the intersection $A \cap B$ is strongly regular.*

Proof. If the intersection is strongly regular the requirements of Theorem 40 are globally fulfilled for A, B subspaces, so AAR converges with linear rate

$$\tilde{c} = \sqrt{1 - \frac{(1-c)}{\kappa^2}} < 1.$$

On the other hand for $\bar{x} \in A \cap B$ by [6, Thm 3.5] we get the characterization

$$\begin{aligned}\text{Fix } T_{AAR} &= (A \cap B) + N_{A-B}(0) \\ &= (A \cap B) + (N_A(\bar{x}) \cap -N_B(\bar{x})).\end{aligned}$$

and so the fix point set of T_{AAR} does not coincide with the intersection if this is not strongly regular. \square

Remark 45 (Friedrichs angle [15]). The Friedrichs angle being less than 1 is not sufficient for convergence of AAR. This can be seen by example 5 (ii). A detailed analysis regarding the relation between the Friedrichs angle and linear convergence of MAP can be found in [13, 14].

4 Concluding Remarks

We state the following theorem to suggest that the framework presented in this work can be extended to a more general setting, for example the adaptive framework discussed in [7]. It shows that the (S, ε) -firm nonexpansiveness is preserved under convex combination of operators.

Theorem 46. *Let T_1 be (S, ε_1) -firmly nonexpansive and T_2 be (S, ε_2) -firmly nonexpansive on D . The convex combination $\lambda T_1 + (1 - \lambda)T_2$ is (S, ε) -firmly nonexpansive on D where $\varepsilon = \max\{\varepsilon_1, \varepsilon_2\}$.*

Proof. Let $x, y \in D$. Let

$$\begin{aligned}x_+ &\in \lambda T_1 x + (1 - \lambda) T_2 x, \quad \text{and} \\ y_+ &\in \lambda T_1 y + (1 - \lambda) T_2 y, \\ \Rightarrow \quad x_+ &= \lambda x_+^{(1)} + (1 - \lambda) x_+^{(2)}, \quad \text{where } x_+^{(1)} \in T_1 x, x_+^{(2)} \in T_2 x \\ y_+ &= \lambda y_+^{(1)} + (1 - \lambda) y_+^{(2)}, \quad \text{where } y_+^{(1)} \in T_1 y, y_+^{(2)} \in T_2 y.\end{aligned}$$

By Lemma 13 (ii) one has nonexpansiveness of the mappings given by $2T_1 x - x$ and $2T_2 x - x$, $x \in D$ that is

$$\begin{aligned}\| [2x_+^{(1)} - x] - [2y_+^{(1)} - y] \| &\leq \sqrt{1 + 2\varepsilon_1} \|x - y\|, \\ \| [2x_+^{(2)} - x] - [2y_+^{(2)} - y] \| &\leq \sqrt{1 + 2\varepsilon_2} \|x - y\|.\end{aligned}$$

This implies

$$\begin{aligned}&\| (2x_+ - x) - (2y_+ - y) \| \\ &= \left\| \left(2 \left[\lambda x_+^{(1)} + (1 - \lambda) x_+^{(2)} \right] - x \right) - \left(2 \left[\lambda y_+^{(1)} + (1 - \lambda) y_+^{(2)} \right] - y \right) \right\| \\ &= \left\| \lambda \left([2x_+^{(1)} - x] - [2y_+^{(1)} - y] \right) - (1 - \lambda) \left([2x_+^{(2)} - x] - [2y_+^{(2)} - y] \right) \right\| \\ &\leq \lambda \left\| [2x_+^{(1)} - x] - [2y_+^{(1)} - y] \right\| + (1 - \lambda) \left\| [2x_+^{(2)} - x] - [2y_+^{(2)} - y] \right\| \\ &\leq \sqrt{1 + 2\varepsilon} \|x - y\|.\end{aligned}$$

Now using Lemma 13 (i) the proof is complete. \square

Since the modulus of linear regularity does not recover optimal convergence results (see Remark 29), the question remains whether there is a quantitative primal definition of an angle between two sets that recovers the same results for MAP as [9]. This could also be useful to achieve optimal linear convergence results for AAR. Another direction for further development is the metric regularity perspective briefly discussed in Proposition 33. This, we believe, will be helpful in extending the current applications beyond projection-type algorithms.

Acknowledgements

The authors gratefully acknowledge the support of DFG-SFB grant 755-TPC2.

References

- [1] F. A. Artacho and J. Borwein. Global convergence of a nonconvex Douglas-Rachford iteration. *J. of Global Optim.*, 2012.
- [2] F. A. Artacho, A. Dontchev, and M. Geoffroy. Convergence of the proximal point method for metrically regular mappings. *ESAIM: Proc.*, 17:1–8, 2007.
- [3] F. A. Artacho and M. Geoffroy. Uniformity and inexact version of a proximal method for metrically regular mappings. *J. Math. Anal. Appl.*, Vol. 335(1):168–183, 2007.
- [4] H. Attouch, J. Bolte, P. Redont, and A. Soubeyran. Proximal alternating minimization and projection methods for nonconvex problems: An approach based on the Kurdyka-Lojasiewicz inequality. *Math. Oper. Res.*, 35(2):438–457, 2010.
- [5] H. Bauschke and P. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, 2011.
- [6] H. Bauschke, P. Combettes, and D. Luke. Finding best approximation pairs relative to two closed convex sets in hilbert spaces. *J. Approx. Theory*, 127:178–314, 2004.
- [7] H. H. Bauschke and J. M. Borwein. On projection algorithms for solving convex feasibility problems. *SIAM Rev.*, 38(3):367–426, 1996.
- [8] H. H. Bauschke, D. R. Luke, H. M. Phan, and X. Wang. Restricted normal cones and sparsity optimization with affine constraints. arXiv:1205.0320, 2012.
- [9] H. H. Bauschke, D. R. Luke, H. M. Phan, and X. Wang. Restricted normal cones and the method of alternating projections. arXiv:1205.0318, 2012.
- [10] J. M. Borwein and B. Sims. The Douglas-Rachford algorithm in the absence of convexity. *Fixed-Point Algorithms for Inverse Problems in Science and Engineering, Springer Optimization and its Applications*, 49:93–109, 2011.
- [11] R. S. Burachik and A. N. Iusem. *Set-Valued Mappings and Enlargements of Monotone Operators*, volume 8 of *Optimization and Its Applications*. Springer, 2008.
- [12] F. Deutsch. *Best Approximation in Inner Product Spaces*. Springer-Verlag, New York, 2001.
- [13] F. Deutsch and H. Hundal. The rate of convergence for the cyclic projections algorithm I: Angles between convex sets. *J. Approx. Theory*, 142(1):36–55, 2006.

- [14] F. Deutsch and H. Hundal. The rate of convergence for the cyclic projections algorithm III: Regularity of convex sets. *J. Approx. Theory*, 155(1):155–184, 2008.
- [15] K. Friedrichs. On certain inequalities and characteristic value problems for analytic functions and for functions of two variables. *Trans. Amer. Math. Soc.*, 41:321–364, 1937.
- [16] K. Goebel and W. A. Kirk. *Topics in Metric Fixed Point Theory*. Cambridge University Press, Cambridge, 1990.
- [17] K. Goebel and S. Reich. *Uniform convexity, hyperbolic geometry, and nonexpansive mappings*. Pure and Applied Mathematics, Vol. 83. Basel: Marcel Dekker, 1984.
- [18] L. G. Gubin, B. T. Polyak, and E. Raik. The method of projections for finding the common point of convex sets. *Comput. Math. Math. Phys.*, 7(6):1–24, 1967.
- [19] A. D. Ioffe. Metric regularity and subdifferential calculus. *Russian Mathematical Surveys*, 55(3):501, 2000.
- [20] A. Iusem, T. Pennanen, and B. Svaiter. Inexact versions of the proximal point algorithm without monotonicity. *SIAM J. Optim.*, 13:1080–1097, 2003.
- [21] D. Klatte and B. Kummer. Optimization methods and stability of inclusions in Banach spaces. 2009.
- [22] A. Y. Kruger. Weak stationarity: Eliminating the gap between necessary and sufficient conditions. *Optimization*, 53:147–164, 2004.
- [23] A. Y. Kruger. About regularity of collections of sets. *Set-Valued Anal.*, 14:187–206, 2006.
- [24] A. Lewis, D. Luke, and J. Malick. Local linear convergence for alternating and averaged nonconvex projections. *Found. Comput. Math.*, 9(4):485–513, 2009.
- [25] B. Mordukhovich. *Variational Analysis and Generalized Differentiation, I: Basic Theory; II: Applications*. Grundlehren der mathematischen Wissenschaften. Springer-Verlag, New York, 2006.
- [26] H. V. Ngai and M. Théra. Metric inequality, subdifferential calculus and applications. *Set-Valued Anal.*, 9:187–216, 2001. 10.1023/A:1011291608129.
- [27] T. Pennanen. Local convergence of the proximal point algorithm and multiplier methods without monotonicity. *Math. Oper. Res.*, 27:170–191, 2002.
- [28] R. A. Poliquin, R. T. Rockafellar, and L. Thibault. Local differentiability of distance functions. *Trans. Amer. Math. Soc.*, 352(11):5231–5249, 2000.
- [29] R. T. Rockafellar and R. J. Wets. *Variational Analysis*. Grundlehren der mathematischen Wissenschaften. Springer-Verlag, Berlin, 1998.
- [30] E. H. Zarantonello. Projections on convex sets in Hilbert space and spectral theory. In E. H. Zarantonello, editor, *Contributions to Nonlinear Functional Analysis*, pages 237–424. Academic Press, New York, 1971.